Standard Libraries

- File i/o
- Regular expression
- Datetime
- Math(numerical and mathematical)

Write, Read,Analysis -- Data Science and Analysis

File HAndling in Python

- File:- Document containing information resides on the permanent stprage
- Different types of files :- txt,doc,pdf,csv and etc ..
- Input --Keyboard
- Output --File ### Modes of the File I/O
- 'w' -- This mode is used to file writing -- If the file is not present first it creates the file and write so me data to it -- If the file is already present then it will rewrite the previous content

In [1]:

```
# Function to create a file and write to the file
def createFile(filename):
    f= open(filename,'w')
    for i in range(10):
        f.write('This is %d Line ' % i)
    print("File is created and data has written")
    return
createFile('file1.txt')
```

```
File is created and data has written
```

In [2]:

```
ls
```

```
 Volume in drive C has no label.
 Volume Serial Number is BC66-BD99

 Directory of C:\Users\mplab

11-07-2019  09:32    <DIR>          .
11-07-2019  09:32    <DIR>          ..
17-06-2019  13:14    <DIR>          .anaconda
05-07-2019  12:58             3,172 .bash_history
21-06-2019  17:16    <DIR>          .conda
21-06-2019  14:30                43 .condarc
29-06-2019  10:37    <DIR>          .idlerc
11-07-2019  09:12    <DIR>          .ipynb_checkpoints
17-06-2019  13:42    <DIR>          .ipython
21-06-2019  14:30    <DIR>          .jupyter
29-06-2019  09:23    <DIR>          .matplotlib
19-06-2019  14:59                69 .minttyrc
21-06-2019  16:54             3,313 21-06-2019.ipynb
24-06-2019  09:11             8,385 22-06-2019.ipynb
24-06-2019  12:00             4,150 24-06-2019(dixita).ipynb
25-06-2019  11:53             6,336 25-06-2019(dixita).ipynb
27-06-2019  10:11            18,419 26th june.ipynb
26-06-2019  11:57            11,491 26th june-Copy1.ipynb
27-06-2019  15:41            13,693 27-06-2019.ipynb
28-06-2019  16:35            15,164 28th june.ipynb
29-06-2019  14:38             5,884 29th june.ipynb
17-06-2019  13:09    <DIR>          Anaconda3
10-07-2019  16:26            11,745 Assignment 2.ipynb
09-07-2019  17:28             2,840 Assignment.ipynb
17-06-2019  10:30    <DIR>          Contacts
10-07-2019  16:02    <DIR>          Desktop
10-07-2019  16:25    <DIR>          Documents
10-07-2019  16:27    <DIR>          Downloads
17-06-2019  10:30    <DIR>          Favorites
11-07-2019  09:33               150 file1.txt
17-06-2019  10:30    <DIR>          Links
17-06-2019  10:30    <DIR>          Music
17-06-2019  10:30    <DIR>          Pictures
04-07-2019  10:28    <DIR>          ProblemslovingProgramming
21-06-2019  16:57             3,313 python (dixita).ipynb
08-07-2019  17:18            30,825 python 1.ipynb
05-07-2019  12:43             3,599 python 1-Copy1.ipynb
09-07-2019  15:54            17,042 Python 3.ipynb
11-07-2019  09:15            20,053 python 4.ipynb
11-07-2019  09:32             1,692 Python5.ipynb
05-07-2019  09:12    <DIR>          pythonprog
17-06-2019  10:30    <DIR>          Saved Games
17-06-2019  10:30    <DIR>          Searches
25-06-2019  15:42             4,146 Untitled.ipynb
05-07-2019  12:38             3,599 Untitled1.ipynb
17-06-2019  10:30    <DIR>          Videos
              23 File(s)        189,123 bytes
              23 Dir(s)  183,207,424,000 bytes free
```

In [3]:

```
def createFile(filename):
    f= open(filename,'w')
    f.write('Testing--\n')
    print("File is created and data has written")
    return
createFile('file1.txt')
```

```
File is created and data has written
```

In [4]:

```
def appendData(filename):
    f=open(filename,'a')
    for i in range(10):
        print("This is %d Line\n" %i)
    print("File created and successfully data written")
    return
appendData('file2.txt')
```

```
This is 0 Line

This is 1 Line

This is 2 Line

This is 3 Line

This is 4 Line

This is 5 Line

This is 6 Line

This is 7 Line

This is 8 Line

This is 9 Line

File created and successfully data written
```

In [5]:

```
def appendData(filename):
    f=open(filename,'a')
    for i in range(10):
        f.write("This is %d Line\n" %i)
    print("File created and successfully data written")
    return
appendData('file2.txt')
```

```
File created and successfully data written
```

In [7]:

```
def appendData(filename):
    f=open(filename,'a')
    f.write("New Line 1 \n")
    f.write("New Line 2 \n")
    print("File created and successfully data written")
    f.close()
    return
appendData('file2.txt')
```

```
File created and successfully data written
```

In [8]:

```
# Function to read of the file
def readFileData(filename):
    f=open(filename,'r')
    if f.mode == 'r':
        x = f.read()
        print(x)
    f.close()
    return
readFileData('file2.txt')
```

```
This is 0 Line
This is 1 Line
This is 2 Line
This is 3 Line
This is 4 Line
This is 5 Line
This is 6 Line
This is 7 Line
This is 8 Line
This is 9 Line
New Line 1
New Line 2
New Line 1
New Line 2
```

In [11]:

```
# Function to read the file
def fileOperations(filename,mode):
    with open(filename,mode) as f:
        if f.mode =='r' :
            data = f.read()
            print(data)
        elif f.mode == 'a':
            f.write('Data to the file')
            print('The data successfully written')
    f.close()
    return
filename = input('Enter the file name')
mode = input('Enter the mode of the file')
fileOperations(filename,mode)
```

```
Enter the file namedata.txt
Enter the mode of the filea
The data successfully written
```

In [16]:

```
# Data Analysis
# Word Count Program
def wordCount(filename,word):
    with open(filename, 'r') as f:
        if f.mode == 'r':
            x= f.read()
            li= x.split()    #It splits the string with whitespace
    cnt= li.count(word)
    return cnt
filename = input('Enter the file name :')
word = input('Enter the word :')  #which word count you need
wordCount(filename,word)
```

```
Enter the file name :file2.txt
Enter the word :This
```

Out[16]:

```
10
```

In [17]:

```
# Character count from the given file
def charCount(filename):
    with open(filename,'r') as f:
        if f.mode == 'r':
            x=f.read()
            li = list(x)
    return len(li)
filename = input('Enter the filename : ')
charCount(filename)
```

```
Enter the filename : file2.txt
```

Out[17]:

```
198
```

In [19]:

```
# Function to find the no. of lines in the i/p file
# i/p -- filename(file2.txt)
# o/p -- No of lines (12)

def countOfLines(filename):
    with open(filename,'r') as f:
        if f.mode == 'r':
            x = f.read()
            li = x.split("\n")
    return len(li)
filename = input('Enter the filename : ')
countOfLines(filename)
```

```
Enter the filename : file2.txt
```

Out[19]:

```
15
```

In [22]:

```
# Function to print the Upper and Lower characters
def caseCount(filename):
    cntUpper = 0
    cntLower = 0
    with open(filename,'r') as f:
        if f.mode =='r':
            x = f.read()
            li = list(x)
    for i in li:
        if i.isupper():
            cntUpper += 1  #cntUpper = cntUpper+1
        elif i.islower():
            cntLower += 1   #cntLower = cntLower+1
    output = 'Upper Case = {0} , Lower Case = {1}'.format(cntUpper,cntLower)
    return output
filename = input('Enter the filename : ')
caseCount(filename)
```

```
Enter the filename : file2.txt
```

Out[22]:

```
'Upper Case = 28 , Lower Case = 100'
```

math, random, os

- os package it contains certain methods which works with os

In [23]:

```
ls
```

```
 Volume in drive C has no label.
 Volume Serial Number is BC66-BD99

 Directory of C:\Users\mplab

11-07-2019  12:28    <DIR>          .
11-07-2019  12:28    <DIR>          ..
17-06-2019  13:14    <DIR>          .anaconda
05-07-2019  12:58             3,172 .bash_history
21-06-2019  17:16    <DIR>          .conda
21-06-2019  14:30                43 .condarc
29-06-2019  10:37    <DIR>          .idlerc
11-07-2019  09:12    <DIR>          .ipynb_checkpoints
17-06-2019  13:42    <DIR>          .ipython
21-06-2019  14:30    <DIR>          .jupyter
29-06-2019  09:23    <DIR>          .matplotlib
19-06-2019  14:59                69 .minttyrc
21-06-2019  16:54             3,313 21-06-2019.ipynb
24-06-2019  09:11             8,385 22-06-2019.ipynb
24-06-2019  12:00             4,150 24-06-2019(dixita).ipynb
25-06-2019  11:53             6,336 25-06-2019(dixita).ipynb
27-06-2019  10:11            18,419 26th june.ipynb
26-06-2019  11:57            11,491 26th june-Copy1.ipynb
27-06-2019  15:41            13,693 27-06-2019.ipynb
28-06-2019  16:35            15,164 28th june.ipynb
29-06-2019  14:38             5,884 29th june.ipynb
17-06-2019  13:09    <DIR>          Anaconda3
10-07-2019  16:26            11,745 Assignment 2.ipynb
09-07-2019  17:28             2,840 Assignment.ipynb
17-06-2019  10:30    <DIR>          Contacts
11-07-2019  10:26                16 data.txt
11-07-2019  12:14    <DIR>          Desktop
10-07-2019  16:25    <DIR>          Documents
10-07-2019  16:27    <DIR>          Downloads
17-06-2019  10:30    <DIR>          Favorites
11-07-2019  09:50                11 file1.txt
11-07-2019  10:10               212 file2.txt
17-06-2019  10:30    <DIR>          Links
17-06-2019  10:30    <DIR>          Music
17-06-2019  10:30    <DIR>          Pictures
04-07-2019  10:28    <DIR>          ProblemslovingProgramming
21-06-2019  16:57             3,313 python (dixita).ipynb
08-07-2019  17:18            30,825 python 1.ipynb
05-07-2019  12:43             3,599 python 1-Copy1.ipynb
09-07-2019  15:54            17,042 Python 3.ipynb
11-07-2019  09:15            20,053 python 4.ipynb
11-07-2019  12:28            13,396 Python5.ipynb
05-07-2019  09:12    <DIR>          pythonprog
17-06-2019  10:30    <DIR>          Saved Games
17-06-2019  10:30    <DIR>          Searches
25-06-2019  15:42             4,146 Untitled.ipynb
05-07-2019  12:38             3,599 Untitled1.ipynb
17-06-2019  10:30    <DIR>          Videos
              25 File(s)        200,916 bytes
              23 Dir(s)  183,209,943,040 bytes free
```

In [27]:

```
cd Desktop/probsolvingprogramming/git
```

```
C:\Users\mplab\Desktop\probsolvingprogramming\git
```

In [28]:

```
ls
```

```
 Volume in drive C has no label.
 Volume Serial Number is BC66-BD99

 Directory of C:\Users\mplab\Desktop\probsolvingprogramming\git

04-07-2019  14:55    <DIR>          .
04-07-2019  14:55    <DIR>          ..
04-07-2019  14:54                15 file1.txt
04-07-2019  14:55    <DIR>          test2
               1 File(s)             15 bytes
               3 Dir(s)  183,209,725,952 bytes free
```

In [47]:

```
cd ..
```

```
C:\Users\mplab\Desktop\probsolvingprogramming
```

In [48]:

```
import os
os.listdir('git/')  # Listdir()  --ls
```

Out[48]:

```
['.git',
 '2019',
 'file1.txt',
 'Single Directory',
 'SingleDirectory',
 'test2',
 'TestFolder']
```

In [49]:

```
li = os.listdir('git/')
for i in li:
    print(i)
```

```
.git
2019
file1.txt
Single Directory
SingleDirectory
test2
TestFolder
```

- Older version Python -- os.listdir()
- New version Python -- os.scandir() and pathlib.Path()

In [50]:

```
li = os.scandir('git/')
for i in li:
    print(i)
```

```
<DirEntry '.git'>
<DirEntry '2019'>
<DirEntry 'file1.txt'>
<DirEntry 'Single Directory'>
<DirEntry 'SingleDirectory'>
<DirEntry 'test2'>
<DirEntry 'TestFolder'>
```

In [51]:

```
from pathlib import Path
li = Path('git/')
for i in li.iterdir():
    print(i.name)
```

```
.git
2019
file1.txt
Single Directory
SingleDirectory
test2
TestFolder
```

Listing all files in a Directory

In [52]:

```
import os
dirPath = "git/"
for i in os.listdir(dirPath):
    if os.path.isfile(os.path.join(dirPath,i)):
        print(i)
```

```
file1.txt
```

In [53]:

```
pwd
```

Out[53]:

```
'C:\\Users\\mplab\\Desktop\\probsolvingprogramming'
```

In [11]:

```
cd git
```

```
C:\Users\mplab\Desktop\probsolvingprogramming\git
```

In [12]:

```
pwd
```

Out[12]:

```
'C:\\Users\\mplab\\Desktop\\probsolvingprogramming\\git'
```

Listing Subdirectories

In [54]:

```
dirPath = 'git/'
for i in os.listdir(dirPath):
    if os.path.isdir(os.path.join(dirPath,i)):
        print(i)
```

```
.git
2019
Single Directory
SingleDirectory
test2
TestFolder
```

In [55]:

```
from pathlib import Path
dirPath = Path('git/')
for i in dirPath.iterdir():
    if i.is_dir():
        print(i.name)
```

```
.git
2019
Single Directory
SingleDirectory
test2
TestFolder
```

In [56]:

```
dirPath = 'Git/'
with os.scandir(dirPath) as f:
    for i in f:
        if i.is_dir():
            print
```

Creating a Single Directory

In [57]:

```
os.mkdir('SingleDirectory')
```

In [58]:

```
import pathlib
p = pathlib.Path('TestFolder')
p.mkdir()
```

In [20]:

```
ls
```

```
 Volume in drive C has no label.
 Volume Serial Number is BC66-BD99

 Directory of C:\Users\mplab\Desktop\probsolvingprogramming\git

11-07-2019  14:46    <DIR>          .
11-07-2019  14:46    <DIR>          ..
04-07-2019  14:54                15 file1.txt
11-07-2019  14:46    <DIR>          Single Directory
11-07-2019  14:45    <DIR>          SingleDirectory
04-07-2019  14:55    <DIR>          test2
               1 File(s)             15 bytes
               5 Dir(s)  184,857,337,856 bytes free
```

Creating multiple Directories

In [28]:

```
import os
os.makedirs('2019/July/11')
```

In [29]:

```
ls
```

```
 Volume in drive C has no label.
 Volume Serial Number is BC66-BD99

 Directory of C:\Users\mplab\Desktop\probsolvingprogramming\git

11-07-2019  15:00    <DIR>          .
11-07-2019  15:00    <DIR>          ..
11-07-2019  15:00    <DIR>          2019
04-07-2019  14:54                15 file1.txt
11-07-2019  14:46    <DIR>          Single Directory
11-07-2019  14:45    <DIR>          SingleDirectory
04-07-2019  14:55    <DIR>          test2
11-07-2019  14:54    <DIR>          TestFolder
               1 File(s)             15 bytes
               7 Dir(s)  184,855,597,056 bytes free
```

In [59]:

```
pwd
```

Out[59]:

```
'C:\\Users\\mplab\\Desktop\\probsolvingprogramming'
```

In [64]:

```
cd probsolvingprogramming
```

```
C:\Users\mplab\Desktop\probsolvingprogramming
```

In [66]:

```
import os
dirPath = 'git/'
for f_name in os.listdir(dirPath):
    if f_name.endswith('.ipynb'):
        print(f_name)
```

Deleting Files and Directories

In [ ]:

```
import os
data_file ='file1.txt'  #Give the path c:\\Users
os.remove(data_file)
```

In [68]:

```
data_dir = 'TestFolder'
os.rmdir(data_dir)
```

In [ ]:

```
import shutil
data_dir = '2019'
shutil.rmtree(data_dir)
```

In [ ]:

```

```